



Finite-horizon LQ control for unknown discrete-time linear systems via extremum seeking [☆]

Paul Frihauf^a, Miroslav Krstic^{a,*}, Tamer Başar^b

^a Department of Mechanical and Aerospace Engineering, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0411, USA

^b Coordinated Science Laboratory, University of Illinois, 1308 West Main Street, Urbana, IL 61801-2307, USA

ARTICLE INFO

Article history:

Received 14 May 2013

Accepted 14 May 2013

Recommended by Alessandro Astolfi

Available online 23 May 2013

Keywords:

Extremum seeking

Linear quadratic regulator

ABSTRACT

We present a non-model based approach for asymptotic, locally exponentially stable attainment of the optimal open-loop control sequence for unknown, discrete-time linear systems with a scalar input, where not even the dimension of the system is known. This control sequence minimizes the finite-time horizon cost function, which is quadratic in the measured output and in the input. We make no assumptions on the stability of the unknown system, but we do assume that the system is reachable. The proposed algorithm employs the multi-variable discrete-time extremum seeking approach to minimize the cost function, extending results established for the scalar discrete-time extremum seeking method. Simulation results show that the Hessian's condition number, used as a measure of the optimization problem's level of difficulty, increases with both the system's level of instability and the length of the finite horizon for a scalar system. Thus, we suggest solving well-conditioned, shorter time horizon optimal control problems to obtain good initial control estimates for longer time horizon problems if possible. We also show that the algorithm accommodates input constraints by employing the projection operator and introduce a Newton-based discrete-time extremum seeking controller, which removes the convergence rate's dependence on the unknown Hessian.

© 2013 European Control Association. Published by Elsevier Ltd. All rights reserved.

1. Introduction

The construction of optimal control signals for dynamical systems is a much-studied area of control theory. For linear systems, an optimal linear feedback law is obtained by solving the well-known Riccati equation. When not all the states are measured, the separation principle allows for the design of optimal observers to obtain state estimates for use in this feedback law. While the states may not be directly known, a model of the system is assumed to be known precisely. In contrast, we study in this paper the optimal control problem for a discrete-time linear system that is entirely *unknown*—only the system's output is known. Our approach utilizes extremum seeking control, which has been the focus of much recent research both in theory [29,7,1,26,25,12,13,24,15] and in applications [28,32,5,19,9,16,17,8].

Recent studies have been dedicated to the optimal control problem for unknown systems, relying mainly on learning and dynamic programming methods. In [21], an observer and neural network are used to approximate a value function, which is used

in a policy iteration scheme to determine an optimal control strategy for an unknown system whose relative degree is known. Reinforcement learning is utilized in [18,31] to achieve optimal control policy for unknown systems. In [11,27], neural networks are used to approximate the unknown nonlinear discrete-time systems so that the optimal control policy can be found offline with approximate dynamic programming. Techniques similar to those presented in this work are used to achieve desired plasma current profiles in a magnetic fusion reactor in [22].

We solve a finite-time horizon, optimal control problem for a completely unknown discrete-time linear system with a scalar input by using a non-model based, extremum seeking controller to attain the open-loop optimal control sequence. Since only the output values of the system can be measured, an optimal feedback control is not possible and the considered cost functions are restricted to quadratic functions in the input and the output, rather than the state. The control values for each step in the time horizon are determined by the extremum seeking controller, which is driven by the value of the cost function. After each iteration of the unknown system, the cost is computed and used to iterate the extremum seeking controller, updating the control sequence. On the average, the controller is driven by the gradient of the cost function, which is driven to zero, causing the control sequence to converge to the optimal open-loop control sequence. While optimal state feedback control would be preferred, its

[☆]This research was made possible by grants from National Science Foundation, DOE, and AFOSR.

* Corresponding author. Tel.: +1 858 822 1374.

E-mail addresses: pfrihauf@ucsd.edu (P. Frihauf), krstic@ucsd.edu (M. Krstic), basar1@illinois.edu (T. Başar).

design requires knowledge of the system model, which we assume is not available.

Our results follow closely and extend those found in [7] to establish convergence for multi-variable, discrete-time extremum seeking controllers. The algorithm is gradient-based so its convergence rate is dependent on the Hessian of the cost function. Convergence of multi-variable, discrete-time extremum seeking schemes with noisy measurements has been shown in [24], but the proof relies on diminishing gains that effectively turn off the algorithm over time. We utilize the two-time scale averaging theory for discrete time systems [2] to prove convergence for a multi-parameter scheme with constant gains, albeit with noise-free measurements, so that the scheme continually seeks to optimize the measured cost, tracking changes in the system. While we prove a local result, non-local results for extremum seeking controllers have been established in [26].

We also introduce and demonstrate in simulation a Newton-based discrete-time extremum seeking controller, whose continuous-time counterpart is derived and analyzed in [15]. The Newton-based scheme allows the algorithm's convergence rate to be user assigned by removing its dependence on the unknown Hessian. Both the gradient-based and Newton-based designs can accommodate input constraints by utilizing a projection operator to solve input constrained optimization problems.

Systems that are amenable to this framework, i.e., unknown or uncertain systems that are highly repeatable with constant initial conditions, have been the focus of a large body of research known as iterative learning control [20,6,30]. Pulsed laser systems, for example, fit the above description. For unknown systems, iterative learning control utilizes finite-time, repetitive experiments to achieve the tracking of a given reference, whereas in this work, our scheme optimizes the system's output according to the computed cost function. Stability of the unknown system is not required (since we consider a finite-horizon problem); however, the optimization problem becomes more difficult as the level of instability of the unknown system increases or as the time horizon increases, which is shown in Section 6.1 with a scalar system example.

This paper is organized as follows: We introduce the problem statement in Section 2 and the gradient-based discrete-time extremum seeking controller in Section 3. The error system is derived in Section 4 and our convergence result is proved in Section 5. In Section 6, simulations for both unconstrained and input constrained optimizations are presented using the gradient-based and Newton-based schemes. We conclude with Section 7.

2. Problem statement

Consider the single-input linear discrete-time system

$$x_{k+1} = A_k x_k + B_k u_k, \quad (1)$$

$$y_k = C_k x_k, \quad (2)$$

where $x_k \in \mathbb{R}^n$, $u_k \in \mathbb{R}$, $y_k \in \mathbb{R}^p$, and A_k, B_k, C_k denote *unknown* matrices of appropriate dimensions at discrete time k . Only the input u_k and the output y_k are treated as known values. Hence, even the state dimension is unknown.

Despite these significant uncertainties, we wish to find the optimal, open-loop control sequence $\{u_k^*\}_{k=0}^{N-1}$ that minimizes the cost function

$$J(u) = \frac{1}{2} y_N^T \tilde{Q}_N y_N + \frac{1}{2} \sum_{k=0}^{N-1} y_k^T \tilde{Q}_k y_k + R_k u_k^2, \quad (3)$$

where $u = [u_0, \dots, u_{N-1}]^T$ and $\tilde{Q}_k, \tilde{Q}_N \geq 0, R_k > 0$ for all $k \in \{0, \dots, N-1\}$. Namely, we want to solve the discrete-time, finite-time horizon

optimal control problem

$$\min_u J(u), \text{ subject to (1)–(2)} \quad (4)$$

with initial condition x_0 . We seek an open-loop solution instead of an optimal state feedback policy since the system is unknown and state information is not available. The cost function (3) can be written in terms of the state x_k as

$$J(u) = \frac{1}{2} x_N^T Q_N x_N + \sum_{k=0}^{N-1} x_k^T Q_k x_k + R_k u_k^2, \quad (5)$$

where $Q_k = C_k^T \tilde{Q}_k C_k, k \in \{0, \dots, N\}$, which is the standard cost function for the LQ optimal control problem with a positive weight on control at each stage. By substituting the system's state trajectory

$$x_k = \Phi_{k,0} x_0 + \sum_{l=0}^{k-1} \Phi_{k,l+1} B_l u_l, \quad (6)$$

$$\Phi_{ij} = \begin{cases} A_{i-1} A_{i-2} \dots A_j, & i > j \\ I, & i = j, \end{cases} \quad (7)$$

the cost function (5) can be written in terms of only the initial state x_0 and the control u_k . In vector form, we have

$$J(u) = \frac{1}{2} x_0^T F x_0 + x_0^T G u + \frac{1}{2} u^T H u, \quad (8)$$

where

$$F = Q_0 + \theta^T \bar{Q} \theta, \quad (9)$$

$$G = \theta^T \bar{Q} \Lambda, \quad (10)$$

$$H = \Lambda^T \bar{Q} \Lambda + \bar{R}, \quad (11)$$

$$\theta = [\phi_{1,0}, \phi_{2,0}, \dots, \phi_{N,0}]^T, \quad (12)$$

$$\bar{Q} = \text{diag}[Q_1, \dots, Q_N], \quad (13)$$

$$\Lambda = \begin{bmatrix} B_0 & 0 & 0 & \dots & 0 \\ \phi_{2,1} B_0 & B_1 & 0 & & \\ \phi_{3,1} B_0 & \phi_{3,2} B_1 & B_2 & & \\ \vdots & & & \ddots & \\ \phi_{N,1} B_0 & \phi_{N,2} B_1 & \phi_{N,3} B_2 & & B_{N-1} \end{bmatrix}, \quad (14)$$

$$\bar{R} = \text{diag}[R_0, \dots, R_{N-1}], \quad (15)$$

and $\text{diag}[\cdot]$ denotes a block diagonal matrix. The Hessian H is positive definite symmetric since it equals the sum of $\Lambda^T \bar{Q} \Lambda \geq 0$ and $\bar{R} > 0$. If the system model were known, the optimal open-loop control values could be found directly by solving the corresponding quadratic program [10,4].

We now make the following assumption for (1):

Assumption 1. The reachability Gramian [23],

$$W_N = \sum_{l=0}^{N-1} \Phi_{N,l+1} B_l B_l^T \Phi_{N,l+1}^T, \quad (16)$$

is positive definite symmetric.

This assumption is a necessary and sufficient condition for the existence of a control sequence to take the system from any initial state x_0 to any final state x_N , provided there are no hard constraints on the control. Hence, the optimal control sequence is not the trivial solution.

3. Optimal control via extremum seeking

The non-model based optimization strategy known as extremum seeking uses sinusoidal perturbations to estimate, and drive

to zero, the gradient of an unknown function. In this case, the unknown function is the cost function (3) and the minimizer is the optimal sequence $\{u_k^*\}_{k=0}^{N-1}$. This minimizer is found by driving the unknown system (1)–(2) with a control sequence $\{u_k(l)\}_{k=0}^{N-1}$ to obtain the cost value $J(l)$ and then iterating the discrete-time extremum seeking controller to produce the sequence $\{u_k(l+1)\}_{k=0}^{N-1}$, where l denotes the l -th iteration of the algorithm.

The extremum seeking controller is given in vector form by

$$\hat{u}(l) = -\frac{\epsilon K}{z-1}[\xi(l)], \quad (17)$$

$$\xi(l) = M(l)\frac{z-1}{z+h}J(l), \quad (18)$$

$$u(l) = \hat{u}(l) + S(l), \quad (19)$$

where $\hat{u} = [\hat{u}_0, \dots, \hat{u}_{N-1}]^T$, ϵ is a small, positive parameter, K is a positive diagonal matrix, and $h \in (0, 1)$. The notation $P(z)q(l)$ is used to denote the signal in the iteration domain that is the output of the transfer function $P(z)$ driven by $q(l)$, where $P(z)$ operates with respect to the iteration domain. The perturbation signals $M(l)$ and $S(l)$, whose notation is chosen to be consistent with [15], are given by

$$M(l) = \left[\frac{2}{a_0} \cos(\omega_0 l - \varphi_0), \dots, \frac{2}{a_{N-1}} \cos(\omega_{N-1} l - \varphi_{N-1}) \right]^T, \quad (20)$$

$$S(l) = [a_0 \cos(\omega_0 l), \dots, a_{N-1} \cos(\omega_{N-1} l)]^T, \quad (21)$$

with $a_k > 0$ and the modulation frequencies given by $\omega_k = b_k \pi$ for all $k \in \{0, \dots, N-1\}$, where $|b_k| \in (0, 1)$ is a rational number and $\omega_k \neq \omega_i$ for all distinct $i, k \in \{0, \dots, N-1\}$. A block diagram of the controller is shown in Fig. 1.

From (18), we see that each scalar control $u_k(l)$ is driven by the scalar input $W(z)J(l)$ multiplied by the sinusoidal perturbation $(2/a_k) \cos(\omega_k l - \varphi_k)$, where $W(z)$ denotes the washout filter $(z-1)/(z+h)$. On the average, this formulation approximates a gradient-based scheme since the distinct modulation frequencies allow each scalar control $u_k(l)$ to estimate its individual contribution to the cost value $J(l)$, which is a function of $u(l) = [u_0(l), u_1(l), \dots, u_{N-1}(l)]$, i.e., at a fixed l , each scalar control is updated by its estimate of the partial derivative $\partial J / \partial u_k(u(l))$.

4. Error system

To analyze the convergence of the system, we formulate its error dynamics. First, we rewrite the cost function (8) using its Taylor series expansion about the optimum $u^* = [u_0^*, \dots, u_{N-1}^*]^T$,

$$J(u) = J(u^*) + \frac{1}{2}(u-u^*)^T H(u-u^*), \quad (22)$$

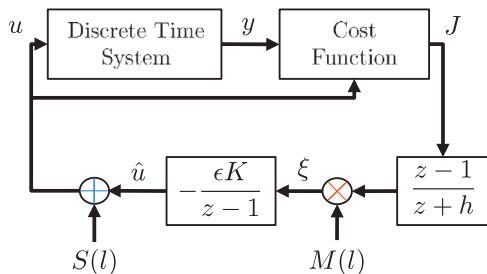


Fig. 1. Discrete-time multi-variable extremum seeking scheme that operates in the iteration domain, indexed by l , to find the optimal control sequence $u^* = \{u_k^*\}_{k=0}^{N-1}$ that minimizes the cost function $J(u)$, where k is the discrete-time index of the unknown system. The depicted discrete-time filters operate with respect to the iteration domain.

and we note that $\nabla J(u^*)$, the gradient of J at u^* , is zero. Next, denote the error relative to the optimal control sequence as

$$\tilde{u}(l) = u(l) - S(l) - u^*, \quad (23)$$

and substitute into (22) to obtain

$$J(\tilde{u}) = J(u^*) + \frac{1}{2} \tilde{u}^T H \tilde{u} + S(l)^T H \tilde{u} + \frac{1}{2} S(l)^T H S(l). \quad (24)$$

Substituting (23) into (17) yields

$$\tilde{u}(l) = -\frac{\epsilon K}{z-1}[\xi(l)] - u^*, \quad (25)$$

which, after noting (18) and (24), can be expressed as the difference equation

$$\begin{aligned} \tilde{u}(l+1) &= \tilde{u}(l) - \epsilon K M(l) W(z) J(u^*) \\ &\quad - \frac{1}{2} \epsilon K M(l) W(z) [\tilde{u}^T H \tilde{u} + S(l)^T H S(l)] \\ &\quad - \epsilon K M(l) W(z) [S(l)^T H \tilde{u}]. \end{aligned} \quad (26)$$

Applying Lemmas 3–5 from the Appendix to the last term of (26) yields, after some algebra

$$\begin{aligned} M(l)W(z)[S(l)^T H \tilde{u}] &= C^-(l) \text{Re}\{\Omega(z, e^{j\omega})[H \tilde{u}]\} \\ &\quad - S^-(l) \text{Im}\{\Omega(z, e^{j\omega})[H \tilde{u}]\} \\ &\quad + C^+(l) \text{Re}\{\Omega(z, e^{j\omega})[H \tilde{u}]\} \\ &\quad - S^+(l) \text{Im}\{\Omega(z, e^{j\omega})[H \tilde{u}]\}, \end{aligned} \quad (27)$$

where, with some notational abuse

$$\Omega(z, e^{j\omega}) \triangleq \text{diag}\{W(e^{j\omega_0} z), \dots, W(e^{j\omega_{N-1}} z)\}, \quad (28)$$

and C^-, S^-, C^+ , and S^+ are $N \times N$ matrices whose k th rows are given by

$$C_k^-(l) \triangleq \left[\frac{a_0}{a_k} \cos((\omega_0 - \omega_k)l + \varphi_k), \dots, \frac{a_{N-1}}{a_k} \cos((\omega_{N-1} - \omega_k)l + \varphi_k) \right] \quad (29)$$

$$S_k^-(l) \triangleq \left[\frac{a_0}{a_k} \sin((\omega_0 - \omega_k)l + \varphi_k), \dots, \frac{a_{N-1}}{a_k} \sin((\omega_{N-1} - \omega_k)l + \varphi_k) \right] \quad (30)$$

$$C_k^+(l) \triangleq \left[\frac{a_0}{a_k} \cos((\omega_0 + \omega_k)l - \varphi_k), \dots, \frac{a_{N-1}}{a_k} \cos((\omega_{N-1} + \omega_k)l - \varphi_k) \right] \quad (31)$$

$$S_k^+(l) \triangleq \left[\frac{a_0}{a_k} \sin((\omega_0 + \omega_k)l - \varphi_k), \dots, \frac{a_{N-1}}{a_k} \sin((\omega_{N-1} + \omega_k)l - \varphi_k) \right]. \quad (32)$$

Note that the diagonal elements of C^- and S^- are time invariant ($\cos(\varphi_k)$ and $\sin(\varphi_k)$, respectively), which will be important for our convergence analysis in Section 5.

To highlight these time invariant terms (27), let C_D^- and S_D^- denote diagonal matrices containing the diagonal elements of C^- and S^- . Then, (27) can be rewritten as

$$\begin{aligned} M(l)W(z)[S(l)^T H \tilde{u}] &= \text{Re}\{\bar{\Omega}(e^{j\omega})\Omega(z, e^{j\omega})[H \tilde{u}]\} \\ &\quad + (C^-(l) - C_D^-) \text{Re}\{\Omega(z, e^{j\omega})[H \tilde{u}]\} \\ &\quad - (S^-(l) - S_D^-) \text{Im}\{\Omega(z, e^{j\omega})[H \tilde{u}]\} \\ &\quad + C^+(l) \text{Re}\{\Omega(z, e^{j\omega})[H \tilde{u}]\} \\ &\quad - S^+(l) \text{Im}\{\Omega(z, e^{j\omega})[H \tilde{u}]\}, \end{aligned} \quad (33)$$

where, again abusing notation, $\bar{\Omega}(e^{j\omega}) = \text{diag}\{e^{j\varphi_0}, \dots, e^{j\varphi_{N-1}}\}$. Substituting (33) into (26) allows the error dynamics to be expressed as follows:

$$\tilde{u}(l+1) - \tilde{u}(l) = \epsilon(L(z)[H \tilde{u}] + \Psi_1^-(l) + \Psi_1^+(l) + \Psi_2(l) + \delta(l)), \quad (34)$$

$$L(z) = -\frac{K}{2}(\bar{\Omega}(e^{j\omega})\Omega(z, e^{j\omega}) + \bar{\Omega}(e^{-j\omega})\Omega(z, e^{-j\omega})), \quad (35)$$

$$\begin{aligned} \Psi_1^-(l) &= K(S^-(l) - S_D^-) \text{Im}\{\Omega(z, e^{j\omega})[H \tilde{u}]\} \\ &\quad - K(C^-(l) - C_D^-) \text{Re}\{\Omega(z, e^{j\omega})[H \tilde{u}]\}, \end{aligned} \quad (36)$$

$$\Psi_1^+(l) = K S^+(l) \text{Im}\{\Omega(z, e^{j\omega})[H \tilde{u}]\} - K C^+(l) \text{Re}\{\Omega(z, e^{j\omega})[H \tilde{u}]\}, \quad (37)$$

$$\Psi_2(l) = -\frac{1}{2}KM(l)W(z)[\tilde{u}^T H\tilde{u}], \tag{38}$$

$$\delta(l) = -\epsilon KM(l)(W(z)[J(u^*)] + \frac{1}{2}W(z)[S(l)^T HS(l)]). \tag{39}$$

In the above formulation, one sees that the error dynamics evolve according to a sum of a linear time-invariant term, $L(z)[H\tilde{u}]$; linear time-varying functions, $\Psi_1^-(l)$ and $\Psi_1^+(l)$; a nonlinear time-varying function $\Psi_2(l)$; and a time-varying function $\delta(l)$ that does not depend on the control error \tilde{u} .

5. Convergence result

To establish the convergence result for the multi-variable, discrete-time extremum seeking scheme, we extend the results found in [7]. First, we prove our main result – local exponential convergence of the homogeneous error system (34) – by employing the two-time scale averaging theory for discrete-time systems [2], and then we consider the full system including the $\delta(l)$ term and establish its convergence.

5.1. Convergence of homogeneous error system

The following theorem states a sufficient condition for locally exponential convergence of the error system (34) without the $\delta(l)$ term. Namely, we have the homogeneous error system that is periodic in time l

$$\tilde{u}(l+1) - \tilde{u}(l) = \epsilon(L(z)[H\tilde{u}] + \Psi_1^-(l) + \Psi_1^+(l) + \Psi_2(l)). \tag{40}$$

Theorem 1. Consider the homogeneous error system (40) with modulation frequencies that satisfy $\omega_k \neq \omega_i$ for all distinct $i, k \in \{0, \dots, N-1\}$ and phase values φ_k selected such that $\text{Re}\{e^{j\varphi_k} W(e^{j\omega_k})\} > 0$ for all $k \in \{0, \dots, N-1\}$. There exists a positive constant ϵ^* such that for all $\epsilon \in (0, \epsilon^*)$, the state-space realization of (40) is locally exponentially stable at the origin.

Proof. Let $(\bar{A}_1, \bar{B}_1, \bar{C}_1, \bar{D}_1)$ be a minimal state space realization of $L(z)$. Since the terms $\Psi_1^-(l)$, $\Psi_1^+(l)$, and $\Psi_2(l)$ have the same structure as the signal in Lemma 6, they can be represented by the minimal state space realizations $(\bar{A}_2, \bar{B}_2, \bar{C}_2(l), \bar{D}_2(l))$, $(\bar{A}_3, \bar{B}_3, \bar{C}_3(l), \bar{D}_3(l))$, and $(\bar{A}_4, \bar{B}_4, \bar{C}_4(l), \bar{D}_4(l))$, respectively. These realizations allow (40) to be written in state-space form as

$$\zeta'(l+1) = \bar{A}\zeta'(l) + h(l, \tilde{u}(l)), \tag{41}$$

$$\tilde{u}(l+1) = \tilde{u}(l) + \epsilon f'(l, \tilde{u}(l), \zeta'(l)), \tag{42}$$

where $\bar{A} = \text{diag}[\bar{A}_1, \bar{A}_2, \bar{A}_3, \bar{A}_4]$,

$$h(l, \tilde{u}(l)) = [\tilde{u}^T H \bar{B}_1^T | \tilde{u}^T H \bar{B}_2^T | \tilde{u}^T H \bar{B}_3^T | \tilde{u}^T H \bar{B}_4^T] (\tilde{u}^T H \tilde{u}) \bar{B}_4^T, \tag{43}$$

$$f'(l, \tilde{u}(l), \zeta'(l)) = \bar{D}_1 H \tilde{u} + \bar{D}_2(l) H \tilde{u} + \bar{D}_3(l) H \tilde{u} + \bar{D}_4(l) (\tilde{u}^T H \tilde{u}) + [\bar{C}_1 | \bar{C}_2(l) | \bar{C}_3(l) | \bar{C}_4(l)] \zeta'(l). \tag{44}$$

The mixed time scale system (41)–(42), with the exponentially stable matrix \bar{A} , can be transformed to permit the application of the two-time scale averaging theory [2]. Let

$$\zeta(l) = \zeta'(l) - w(l, \tilde{u}), \tag{45}$$

$$w(l, \tilde{u}) = \sum_{i=0}^{l-1} \bar{A}^{l-i-1} h(i, \tilde{u}), \tag{46}$$

with $h(l, \tilde{u})$ shown in (43). The transformed system is then

$$\zeta(l+1) = \bar{A}\zeta(l) + \epsilon g(l, \tilde{u}, \zeta), \tag{47}$$

$$\tilde{u}(l+1) = \tilde{u}(l) + \epsilon f(l, \tilde{u}, \zeta), \tag{48}$$

where

$$\begin{aligned} \epsilon g(l, \tilde{u}, \zeta) &= \bar{A}w(l, \tilde{u}(l)) + h(l, \tilde{u}(l)) - w(l+1, \tilde{u}(l+1)), \\ &= -\left(\int_0^1 \frac{\partial w}{\partial \tilde{u}}(l+1, s\tilde{u}(l+1) + (1-s)\tilde{u}(l)) ds \right) \\ &\quad \epsilon f'(l, \tilde{u}, \zeta + w(l, \tilde{u})), \end{aligned} \tag{49}$$

$$f(l, \tilde{u}, \zeta) = f'(l, \tilde{u}(l), \zeta + w(l, \tilde{u})). \tag{50}$$

To prove convergence, we analyze the averaged system of (48), given by

$$\tilde{u}^{\text{ave}}(l+1) = \tilde{u}^{\text{ave}}(l) + \epsilon f^{\text{ave}}(\tilde{u}^{\text{ave}}(l)), \tag{51}$$

where f^{ave} is computed using the operator $\mathcal{A}\{\cdot\}$ [2],

$$f^{\text{ave}}(\tilde{u}) = \mathcal{A}\{f(l, \tilde{u}, 0)\} = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{l=s+1}^{s+T} f(l, \tilde{u}, 0). \tag{52}$$

Rather than attempting to compute $f^{\text{ave}}(\tilde{u})$ directly from (52), we use (44)–(46), and (50) to obtain

$$\begin{aligned} f(l, \tilde{u}, 0) &= f'(l, \tilde{u}, w(l, \tilde{u})), \\ &= \bar{D}_1 H \tilde{u} + \bar{D}_2(l) H \tilde{u} + \bar{D}_3(l) H \tilde{u} + \bar{D}_4(l) (\tilde{u}^T H \tilde{u}) \\ &\quad + [\bar{C}_1 | \bar{C}_2(l) | \bar{C}_3(l) | \bar{C}_4(l)] \\ &\quad \times \sum_{i=0}^{l-1} \bar{A}^{l-i-1} [\tilde{u}^T H \bar{B}_1^T | \tilde{u}^T H \bar{B}_2^T | \tilde{u}^T H \bar{B}_3^T | \tilde{u}^T H \bar{B}_4^T] (\tilde{u}^T H \tilde{u}) \bar{B}_4^T, \\ &= L(z)[H\tilde{u}] + \Psi_1^-(l) + \Psi_1^+(l) + \Psi_2(l), \end{aligned} \tag{53}$$

where \tilde{u} is considered a constant, and thus,

$$f^{\text{ave}}(\tilde{u}) = \mathcal{A}\{L(z)[H\tilde{u}] + \Psi_1^-(l) + \Psi_1^+(l) + \Psi_2(l)\}. \tag{54}$$

Computing the individual terms yields

$$\begin{aligned} \mathcal{A}\{\Psi_1^-(l)\} &= \mathcal{A}\{K(S^-(l) - S_D^-) \text{Im}\{\Omega(z, e^{j\omega})\} [H\tilde{u}]\} \\ &\quad - K(C^-(l) - C_D^-) \text{Re}\{\Omega(z, e^{j\omega})\} [H\tilde{u}]\}, \\ &= 0, \end{aligned} \tag{55}$$

$$\begin{aligned} \mathcal{A}\{\Psi_1^+(l)\} &= \mathcal{A}\{KS^+(l) \text{Im}\{\Omega(z, e^{j\omega})\} [H\tilde{u}]\} \\ &\quad - KC^+(l) \text{Re}\{\Omega(z, e^{j\omega})\} [H\tilde{u}]\}, \\ &= 0, \end{aligned} \tag{56}$$

$$\begin{aligned} \mathcal{A}\{\Psi_2(l)\} &= \mathcal{A}\{-KM(l)W(z)[\tilde{u}^T H\tilde{u}]\}, \\ &= 0, \end{aligned} \tag{57}$$

where, again, $\tilde{u}(l)$ is considered a constant sequence and we have noted that $\omega_k \neq \omega_i$ for all $i \neq k$. The final term, which yields the averaged system, is

$$\begin{aligned} f^{\text{ave}}(\tilde{u}) &= \mathcal{A}\{L(z)[H\tilde{u}]\}, \\ &= \mathcal{A}\{-\frac{1}{2}K(\bar{\Omega}(e^{j\varphi})\Omega(z, e^{j\omega}) + \bar{\Omega}(e^{-j\varphi})\Omega(z, e^{-j\omega})) [H\tilde{u}]\}, \\ &= -K\Sigma H\tilde{u}, \end{aligned} \tag{58}$$

where $\Sigma > 0$ is a diagonal matrix with elements $\sigma_k = \text{Re}\{e^{j\varphi_k} W(e^{j\omega_k})\} = |W(e^{j\omega_k})| \times \cos(\psi_k + \varphi_k)$, $\psi_k = \angle W(e^{j\omega_k})$, and $W(e^{j\omega_k}) = (e^{j\omega_k} - 1)/(e^{j\omega_k} + h)$. Hence, from (51) and (58), the averaged system is

$$\tilde{u}^{\text{ave}}(l+1) = (I - \epsilon K\Sigma H)\tilde{u}^{\text{ave}}(l), \tag{59}$$

where I denotes the $N \times N$ identity matrix.

Let $V(\tilde{u}^{\text{ave}}) = (\tilde{u}^{\text{ave}})^T \bar{K}^{-1} \tilde{u}^{\text{ave}}$ be a Lyapunov function where $\bar{K} = K\Sigma$. Then, computing the difference $\Delta V(\tilde{u}^{\text{ave}}) = V(\tilde{u}^{\text{ave}}(l+1)) - V(\tilde{u}^{\text{ave}}(l))$ yields

$$\begin{aligned} \Delta V(\tilde{u}^{\text{ave}}) &= (\tilde{u}^{\text{ave}})^T (I - \epsilon \bar{K}H)^T \bar{K}^{-1} (I - \epsilon \bar{K}H)\tilde{u}^{\text{ave}} - (\tilde{u}^{\text{ave}})^T \bar{K}^{-1} \tilde{u}^{\text{ave}}, \\ &= (\tilde{u}^{\text{ave}})^T (-2\epsilon H + \epsilon^2 H\bar{K}H)\tilde{u}^{\text{ave}}, \\ &\leq -2\epsilon \lambda_{\min}(H) |\tilde{u}^{\text{ave}}|^2 + \epsilon^2 \lambda_{\max}(H\bar{K}H) |\tilde{u}^{\text{ave}}|^2, \end{aligned} \tag{60}$$

where $|\cdot|$ denotes the Euclidean norm. Selecting $\epsilon = \lambda_{\min}(H)/\lambda_{\max}(H\bar{K}H)$ results in

$$\Delta V(\tilde{u}^{\text{ave}}) \leq -\frac{(\lambda_{\min}(H))^2}{\lambda_{\max}(H\bar{K}H)} |\tilde{u}^{\text{ave}}|^2. \quad (61)$$

For all $\epsilon \in (0, 2\lambda_{\min}(H)/\lambda_{\max}(H\bar{K}H))$, the averaged system is exponentially stable, completing the proof [2, Theorem 2.4]. \square

The eigenvalues of H , denoted by $\lambda(H)$, play a prominent role in the convergence rate of the extremum seeking controller, and since H equals the sum of a positive semidefinite symmetric matrix and a positive definite symmetric matrix – a topic that has been the focus of much research [3,14] – we can develop bounds on $\lambda(H)$.

Lemma 1. Let $\beta_1 \geq \beta_2 \geq \dots \geq \beta_N$ be the ordered eigenvalues of H . Similarly, let $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$ and $\rho_1 \geq \rho_2 \geq \dots \geq \rho_N$ be the ordered eigenvalues of $\Lambda^T \bar{Q} \Lambda$ and \bar{R} , respectively. Then,

$$\max_{i+j=N+1} \{\lambda_i + \rho_j\} \leq \beta_1 \leq \min_{i+j=l+1} \{\lambda_i + \rho_j\}. \quad (62)$$

If the input penalty is constant, i.e., $R_k \equiv R$, then

$$\beta_1 = \lambda_1 + R. \quad (63)$$

Proof. Since $H = \Lambda^T \bar{Q} \Lambda + \bar{R}$, where $\Lambda^T \bar{Q} \Lambda \geq 0$ and $\bar{R} > 0$, (62) is immediate [3,14]. If the input penalty is constant, then $\rho_j = R$ for all j and

$$\max_{i+j=N+1} \{\lambda_i + R\} = \lambda_1 + R \leq \beta_1, \quad (64)$$

$$\min_{i+j=l+1} \{\lambda_i + R\} = \lambda_l + R \geq \beta_l, \quad (65)$$

which proves (63). \square

It is important to note that $\lambda(H)$ is bounded below by the smallest input weight $\min_k R_k$, which is a strictly positive value. Since our system is unknown and possibly high dimensional, the matrix $\Lambda^T \bar{Q} \Lambda$ is likely to have some, if not many, zero eigenvalues due to the sparsity of \bar{Q} , which means a subset of $\{\beta_i\}_{i=1}^N$ will be precisely equal to R when the input penalty is constant.

If the gain K and phase values φ_i are chosen so that the elements of $K\Sigma$ in (59) all equal $\nu > 0$, we can rewrite (59) as

$$\tilde{u}^{\text{ave}}(l+1) = r \tilde{u}^{\text{ave}}(l) \quad (66)$$

where $r = (I - \epsilon \nu H)$, and characterize the eigenvalues of r .

Corollary 1. Consider the system (66) with $r = (I - \epsilon \nu H)$, $H > 0$, and scalars $\epsilon, \nu > 0$. The eigenvalues of r are given by

$$\lambda_i(r) = 1 - \epsilon \nu \lambda_i(H), \quad i \in \{1, \dots, N\}, \quad (67)$$

with bounds on $\lambda(H)$ given in Lemma 1.

Proof. Let $H = UTU^T$ be the Schur decomposition of H , where T is upper triangular and U is unitary. Then

$$\begin{aligned} U^T r U &= U^T (I - \epsilon \nu H) U, \\ &= I - \epsilon \nu T, \end{aligned} \quad (68)$$

which is upper triangular with diagonal elements equal to (67). \square

5.2. Convergence of full error system

With the exponential stability of the averaged homogeneous error system established, we now consider the full system (34). First, we state the convergence properties of $\delta(l)$ in the following lemma:

Lemma 2. The time-varying function $\delta(l)$ exponentially converges to an $O(\epsilon|a|)$ -neighborhood of zero:

$$|\delta(l)| \leq \epsilon^{-l} + c_1 \epsilon |a|, \quad (69)$$

where c_1 is a constant and $a = [a_0, \dots, a_{N-1}]$.

Proof. Rewrite (39) as $\delta(l) = \delta_1(l) + \delta_2(l)$, where

$$\delta_1(l) = -\epsilon KM(l)W(z)[U^*], \quad (70)$$

$$\delta_2(l) = -\frac{\epsilon}{2} KM(l)W(z)[S(l)^T HS(l)]. \quad (71)$$

The washout filter $W(z)$ has zero DC gain, so $\delta_1(l)$ has only exponentially decaying terms. The remainder of the proof consists of bounding $\delta_2(l)$, which requires bounding $M(l)$ and $W(z)[S(l)^T HS(l)]$.

The norm of the perturbation signals $M(l)$ can be bounded as follows:

$$|M(l)|^2 = \sum_{i=0}^{N-1} \frac{4}{a_i^2} \cos^2(\omega_i l - \varphi_i) \leq \sum_{i=0}^{N-1} \frac{4}{a_i^2} \leq \frac{4N}{\min\{a_i^2\}}. \quad (72)$$

For $W(z)[S(l)^T HS(l)]$, we begin by rewriting $S(l)^T HS(l)$ as a double summation

$$|W(z)[S(l)^T HS(l)]|^2 = \left| W(z) \left[\sum_{i=1}^N \sum_{j=1}^N h_{ij} a_i a_j \cos(\omega_i l) \cos(\omega_j l) \right] \right|^2, \quad (73)$$

where h_{ij} denotes the ij -th element of H . Applying the linearity of $W(z)$ and the trigonometric identity $\cos(\theta) \cos(\psi) = (1/2) \cos(\theta - \psi) + (1/2) \cos(\theta + \psi)$ yields

$$\begin{aligned} |W(z)[S(l)^T HS(l)]|^2 &= \frac{1}{4} \left| \sum_{i=1}^N \sum_{j=1}^N h_{ij} a_i a_j W(z) [\cos((\omega_i - \omega_j)l) + \cos((\omega_i + \omega_j)l)] \right|^2, \\ &= \frac{1}{4} \left| \sum_{i=1}^N \sum_{j=1}^N h_{ij} a_i a_j \left(|W(e^{j(\omega_i - \omega_j)})| \cos((\omega_i - \omega_j)l + \psi_{ij}^-) \right. \right. \\ &\quad \left. \left. + |W(e^{j(\omega_i + \omega_j)})| \cos((\omega_i + \omega_j)l + \psi_{ij}^+) \right) \right|^2, \end{aligned} \quad (74)$$

where $\psi_{ij}^- = \angle W(e^{j(\omega_i - \omega_j)})$ and $\psi_{ij}^+ = \angle W(e^{j(\omega_i + \omega_j)})$. Let $\mathcal{W} = \{\omega_i - \omega_j, \omega_i + \omega_j\}$ for all distinct $i, j \in \{0, \dots, N-1\}$, then (74) has the bound

$$\begin{aligned} |W(z)[S(l)^T HS(l)]|^2 &\leq \frac{1}{4} \max_{\omega \in \mathcal{W}} |W(e^{j\omega})|^2 \\ &\quad \left| \sum_{i=1}^N \sum_{j=1}^N h_{ij} a_i a_j (\cos((\omega_i - \omega_j)l + \psi_{ij}^-) \right. \\ &\quad \left. + \cos((\omega_i + \omega_j)l + \psi_{ij}^+)) \right|^2, \\ &\leq \max_{\omega \in \mathcal{W}} |W(e^{j\omega})|^2 |a^T H a|^2, \\ &\leq \max_{\omega \in \mathcal{W}} |W(e^{j\omega})|^2 \lambda_{\max}(H) |a|^2. \end{aligned} \quad (75)$$

From (71), (72), and (75), we have

$$|\delta_2(l)|^2 \leq N \lambda_{\max}(K) \lambda_{\max}(H) \max_{\omega \in \mathcal{W}} |W(e^{j\omega})|^2 \epsilon^2 \frac{|a|^2}{\min\{a_i^2\}}, \quad (76)$$

which completes the proof. \square

From the perturbed averaged system

$$\tilde{u}^{\text{ave}}(l+1) = (I - \epsilon \bar{K} H) \tilde{u}^{\text{ave}}(l) + \delta(l), \quad (77)$$

and Lemma 2, we see that $\tilde{u}^{\text{ave}}(l)$ converges exponentially to an $O(|a|)$ -neighborhood of the origin since $|\delta_k(l)| \leq \epsilon^{-l} + c_1 \epsilon |a|$. From [2], the exponential convergence rate of \tilde{u} in (34) tends to the rate of \tilde{u}^{ave} in the average system as ϵ goes to zero. We can now state the convergence result for the overall \tilde{u} -system.

Theorem 2. Consider the full system (34) with the conditions of Theorem 1 satisfied. For sufficiently small $a_k, k \in \{0, \dots, N-1\}$, there exists $\epsilon_1^* \in (0, \epsilon^*]$, such that for all $\epsilon \in (0, \epsilon_1^*]$, the error variable \tilde{u} locally exponentially converges to an $O(|a|)$ -neighborhood of the origin.

Corollary 2. With the conditions of Theorem 2 satisfied, the cost value J locally exponentially converges to an $O(|a|^2)$ -neighborhood of the optimal cost $J(u^*)$.

Proof. Define $\tilde{J}(u) = J(u) - J(u^*)$. Then, from (22) and (24), we have
$$\tilde{J}(u) = \frac{1}{2}(u - u^*)^T H(u - u^*),$$

$$= \frac{1}{2} \tilde{u}^T H \tilde{u} + S(l)^T H \tilde{u} + \frac{1}{2} S(l)^T H S(l).$$
 (78)

From Theorem 2, \tilde{u} locally exponentially converges to an $O(|a|)$ -neighborhood of the origin. Thus, $\tilde{J}(u)$ locally exponentially converges to an $O(|a|^2)$ -neighborhood of the origin. \square

6. Simulation results

We first present a scalar example to explore how the instability of the discrete-time system affects the optimization problem's structure before demonstrating the extremum seeking controller in simulation with an unstable third-order system. The example solves both an unconstrained and an input-constrained optimal control problem. Lastly, a Newton-based discrete-time extremum seeking design is introduced and simulated.

6.1. Scalar system example

Consider the scalar system

$$x_{k+1} = \alpha x_k + u_k, \tag{79}$$

with initial condition $x_0 = 1$ and the cost function $J = \frac{1}{2} x_N^2 + \frac{1}{2} \sum_{k=0}^{N-1} x_k^2 + u_k^2$, where $Q_N = Q_k = R_k = 1$. To study how the stability of (79) affects the eigenvalues of H , we will vary the parameter α in the interval $[0, 3]$ for intervals $N \in \{2, 3, 5, 8, 10\}$.

Fig. 2(a) depicts how the condition number of the Hessian, $\kappa(H)$, varies with α for each time horizon N . The curves indicate that the optimization problem becomes more difficult as the instability of the discrete-time system increases and this difficulty is more acute for longer time horizons. In fact, $\kappa(H)$ appears to grow almost exponentially with α with a rate of increase dependent on the time horizon N . Fig. 2(b) shows how, for this example, the eigenvalues appear to grow and spread apart as α increases from 0 to 1. For $\alpha > 1$ the maximal eigenvalue continues to increase while the other eigenvalues converge to $\lambda = 1$, and consequently, the condition number also increases since for the symmetric matrix $H, \kappa(H) = \lambda_{\max}(H) / \lambda_{\min}(H)$.

One should note that these structural changes to the Hessian are due to the inherent challenges of this optimization problem; they are not caused by the extremum seeking algorithm. Thus, if possible, one should consider attaining the optimal sequence $\{u_k^*\}_{k=0}^{N-1}$ recursively to find better initial control sequences for longer time horizons since $\kappa(H)$ grows with N . Specifically, first obtain $\{u_k^*\}_{k=0}^{M-1}$ and use $\{u_0^*, u_1^*, \dots, u_{M-1}^*, u_M\}$ as the initial control sequence for the $M + 1$ finite horizon, where two possible values for u_M are $u_M = 0$ or $u_M = u_{M-1}^*$. Repeat this process until the desired N intervals have been optimized. In practice, the current non-optimal control sequence may be a good initial control sequence for the optimization scheme as well.

6.2. Unstable third-order system example

Consider

$$x_{k+1} = \begin{bmatrix} 0.5 & 0.5 & 0.3 \\ 0.6 & 0.5 & -0.2 \\ 0.5 & 0 & 0.5 \end{bmatrix} x_k + \begin{bmatrix} -0.15 \\ 0.5 \\ 0.75 \end{bmatrix} u_k, \tag{80}$$

$$y_k = \begin{bmatrix} 1 & 0.5 & 0 \\ 0.15 & -1.5 & 0 \end{bmatrix} x_k, \tag{81}$$

with initial condition $x_0 = [-1, -0.5, 0.1]^T$. The system's eigenvalues are $-0.22, 1.11,$ and 0.61 . Also note that the third state does not directly appear in the system's output.

We want to minimize the cost function, $J = \frac{1}{2} \gamma_N y_N^2 + \frac{1}{2} \sum_{k=0}^{N-1} \gamma_k y_k^2 + 5u_k^2$, with penalty weights γ_k , when u_k is unconstrained and when we have input constraints $\underline{u}_k \leq u_k \leq \bar{u}_k$. We denote the lower and upper control limits in vector form as \underline{u} and \bar{u} . For the unconstrained case, we implement the scheme (17)–(19), with $a_k = 0.2, K = -\text{diag}[1.0, 1.25, 1.67, 2.5, 5.0] \times 10^{-4}, \omega_k$ drawn from a uniform distribution such that $\omega_k \in (0, \pi)$, and $\varphi_k = -\angle W(e^{j\omega_k})$. For the constrained case, we use the same parameters but modify the extremum seeking controller to accommodate the input constraints. Specifically, the k -th control loop is implemented as $\hat{u}_k(l + 1) = \hat{u}_k(l) - \text{Proj}\{cK_k \xi(l); \underline{u}_k, \bar{u}_k\}$, where

$$\text{Proj}\{\phi; \underline{u}_k, \bar{u}_k\} = \begin{cases} 0 & \text{if } \hat{u}_k \leq \underline{u}_k + \phi, \\ 0 & \text{if } \hat{u}_k \geq \bar{u}_k + \phi, \\ \phi & \text{otherwise,} \end{cases} \tag{82}$$

is the projection operator used to constrain \hat{u}_k according to the input constraints and K_k is the k -th diagonal element of the gain matrix K . (The input value u_k is allowed to violate the constraints by the perturbation magnitude a_k .) In this example, we study the input bounds $|u_k| \leq 1$.

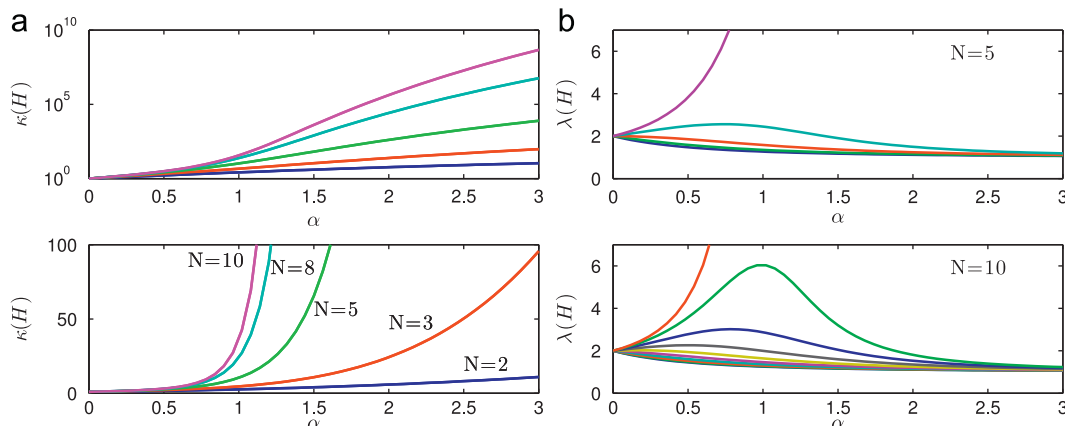


Fig. 2. (a) Condition number of $\kappa(H)$ versus α for various time horizons N with a log scale (top) and a linear scale (bottom) for the y-axis, and (b) eigenvalues of H versus α for $N=5$ (top) and $N=10$ (bottom). The maximum eigenvalue is approximately the condition number $\kappa(H) = \lambda_{\max}(H) / \lambda_{\min}(H)$.

For both the unconstrained and constrained scenarios, the system is simulated with a zero initial control sequence for the finite horizon $N=5$ and penalty weights $\gamma_0=3, \gamma_1=5, \gamma_2=7, \gamma_3=9, \gamma_4=11, \gamma_5=1$. Fig. 3 depicts the planar trajectories of \hat{u}_0 and \hat{u}_1 for both cases with the level sets of $J(u)$ and the input constraints (dashed lines) superimposed. The constrained trajectory hits the upper limit of \hat{u}_0 and climbs along the boundary until reaching the constrained optimum denoted by the cyan square. The optimal control values for the constrained problem are obtained by formulating a quadratic program [10] and minimizing $J(u)$ subject to the input constraints. Fig. 4 depicts the time history of \hat{u} for both cases. Note how the trajectory of \hat{u}_1 is affected by the \hat{u}_0 hitting its constraint in Fig. 4(b).

6.3. Newton-based discrete-time extremum seeking

The extremum seeking scheme introduced in Section 3 and subsequently analyzed is a gradient-based design, and consequently, its convergence rate is governed by the unknown Hessian as seen in (59). The Newton-based discrete-time extremum seeking design, depicted in Fig. 5, utilizes a perturbation matrix $N(l)$ to generate an estimate \hat{H} of the Hessian on the average and a

discrete-time Riccati equation to generate an estimate of the Hessian's inverse, which avoids invertibility issues that may arise by attempting to directly invert \hat{H} since it may be singular or ill-conditioned. (See [15] for derivation and thorough analysis of the continuous-time Newton-based design.)

The Newton-based controller is constructed as follows:

$$\hat{u}(l) = -\frac{\epsilon K \Gamma}{z-1} [\xi(l)], \tag{83}$$

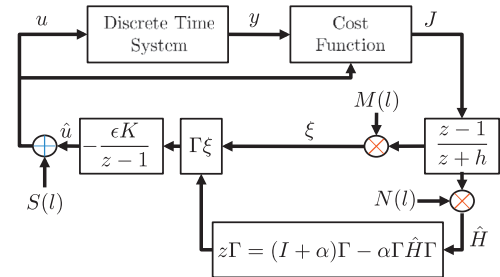


Fig. 5. Newton-based discrete-time extremum seeking scheme to find the optimal control sequence $u^* = \{u_k^*\}_{k=0}^{N-1}$ that minimizes the cost function $J(u)$. The terms ξ and \hat{H} are estimates of the gradient and the Hessian, respectively.

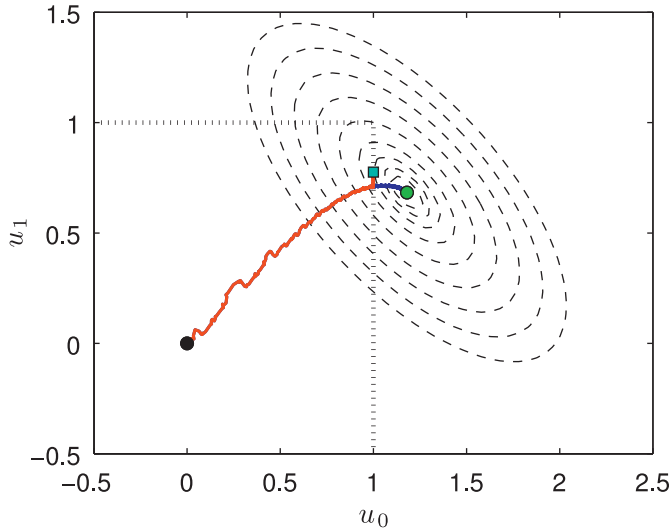


Fig. 3. Trajectory of u_0 and u_1 for unconstrained (blue) and constrained (red) control cases for the unstable third-order system. The cost level sets are shown with the unconstrained minimum (green circle). Dashed lines denote the input constrained set with the optimum value on the boundary (cyan square). (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

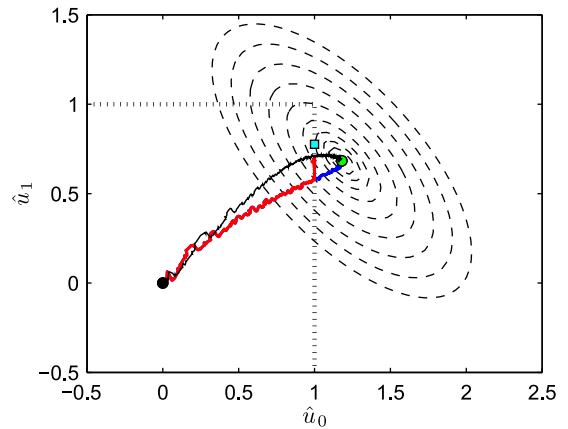


Fig. 6. Trajectory of u_0 and u_1 for unconstrained (blue) and constrained (red) control cases for the unstable third-order system using the Newton-based controller. The cost level sets are shown with the unconstrained minimum (green circle). Dashed lines denote the input constrained set with the optimum value on the boundary (cyan square). For reference, the trajectory generated by the gradient-based controller is also shown (black dashed). (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

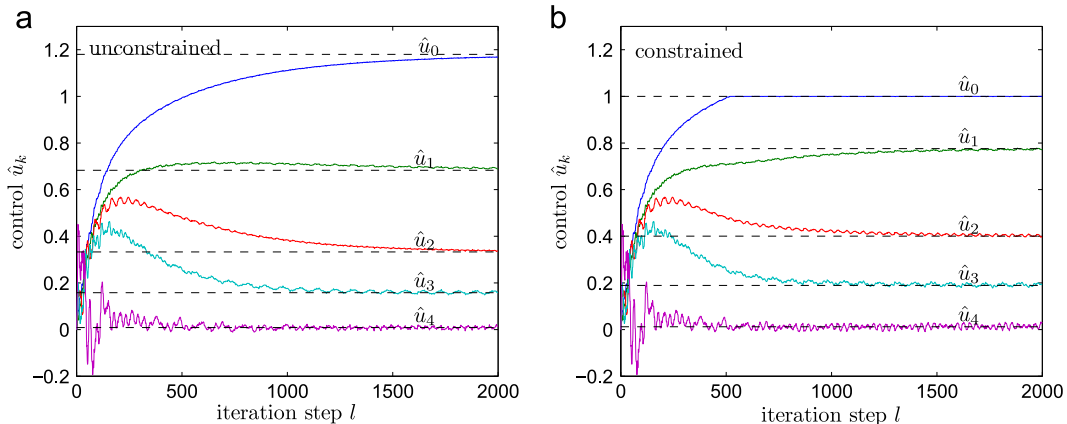


Fig. 4. Time history of \hat{u} for (a) unconstrained and (b) input constrained optimizations.

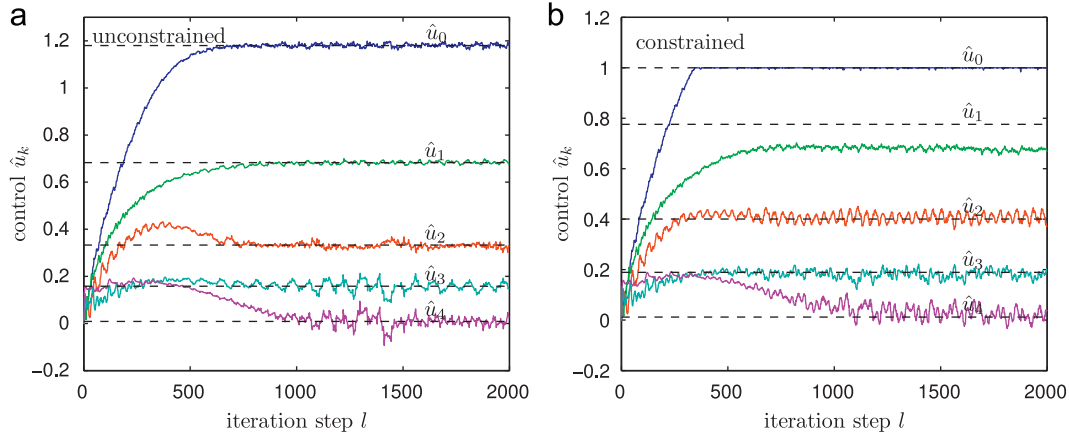


Fig. 7. Time history of \hat{u} for (a) unconstrained and (b) input constrained optimizations when using the Newton-based extremum seeking controller.

$$z\Gamma(l) = (I + \alpha)\Gamma - \alpha\Gamma(N(l)W(z)[J(l)]\Gamma), \quad (84)$$

$$\xi(l) = M(l)W(z)[J(l)], \quad (85)$$

$$u(l) = \hat{u}(l) + S(l), \quad (86)$$

where α is a positive real number, and as in (17)–(19), ϵ is a small, positive parameter, K is a positive diagonal matrix, $h \in (0, 1)$, and $M(l)$ and $S(l)$ are defined by (20) and (21). The $N \times N$ symmetric matrix $N(l)$ consists of the elements

$$N_{i,i}(l) = \frac{16}{a_i^2} \left(\cos^2(\omega_i l) - \frac{1}{2} \right) \quad (87)$$

$$N_{i,j}(l) = \frac{4}{a_i a_j} \cos(\omega_i l) \cos(\omega_j l), \quad i \neq j, \quad (88)$$

whose derivation imposes the frequency requirements

$$\omega_i \notin \{\omega_j, \frac{1}{2}(\omega_j + \omega_k), \omega_j + 2\omega_k, \omega_j + \omega_k \pm \omega_l\} \quad (89)$$

for all distinct $i, j, k, l \in \{0, \dots, N-1\}$. On the average, $N(l)W(z)[J(l)]$ estimates the Hessian H , which allows Γ to converge to H^{-1} .

Following the derivation of the average error system (59), the homogeneous average error system for the Newton-based design can be shown to be

$$\tilde{u} = (I - \epsilon K \Sigma) \tilde{u} - \epsilon K \Sigma \tilde{H} \tilde{u}, \quad (90)$$

$$\tilde{r} = (1 - \alpha) \tilde{r} - \alpha \tilde{r} \Sigma H \tilde{r}, \quad (91)$$

where $\tilde{r} = \Gamma - H^{-1}$. The linearization of this system has its eigenvalues at $(I - \epsilon K \Sigma)$ and $(1 - \alpha)$, which are user assigned and thus, the convergence rate dependence on H has been removed. However, convergence is still local and an initial estimate \tilde{H} that is sufficiently close to H is required, which may be difficult for unstable systems and/or long time horizons based on the example in Section 6.1.

For comparison with the gradient-based simulations, we repeat the optimization problem in Section 6.2 with the Newton-based controller (83)–(86) using the same common parameter values and the new parameters as follows: $\alpha = 0.005$, the elements of K to all be equal to -0.0005 , and $\Gamma(0) = \text{diag}[0.2, 0.25, 0.33, 0.5, 1.0]$. The selection of K and Γ was done such that the initial gain $K\Gamma(0)$ is the same as the gain K used in the gradient-based design.

Fig. 6 depicts the planar trajectories of \hat{u}_0 and \hat{u}_1 for the unconstrained and input constrained cases with the level sets of $J(u)$ and the input constraints (dashed lines) superimposed. Also shown is the trajectory generated by the gradient-based controller, which highlights how the Newton-based method moves directly toward the optimal point. Fig. 7 depicts the time history of \hat{u} for both cases. Note how the rise time for \hat{u}_0 is faster with the Newton-based scheme, compared to the trajectories in Fig. 4. Also

of interest is how in the constrained case \hat{u}_1 does not reach its optimal value, possibly due to the influence of a constrained control value on the estimate of the Hessian.

7. Conclusions

We have introduced a non-model based approach to solve the finite-time horizon optimal control problem for *unknown* discrete-time systems and established its convergence to the open-loop control sequence $\{u_k^*\}_{k=0}^{N-1}$ that minimizes a cost function that is quadratic in the input and output. This result extends the convergence results found in [7]. Such a framework may be a natural tool for the real-time optimization of highly repetitive systems that are unknown and potentially high-dimensional.

Convergence does not depend on the time horizon length N or on the stability of the unknown system (because the time horizon is finite). However, the optimization problem, independent of the solver, becomes more difficult as N increases and as the instability of the system increases. If possible, sequences for shorter time horizons could be found to obtain better initial estimates for solving more difficult longer time horizon optimization problems.

We have also introduced a Newton-based discrete-time extremum seeking controller, whose continuous-time counterpart is developed in [15]. A simulation example shows how the Newton-based scheme takes a more direct trajectory to the optimal point than does the gradient-based scheme. Extending this work to nonlinear systems is of interest.

Appendix A

Lemma 3. For the transfer function $G(z)$ the following is true for any real φ :

$$G(z)[\cos(\omega k - \varphi)v(k)] = \text{Re}\{e^{j(\omega k - \varphi)}G(e^{j\omega z})[v(k)]\}. \quad (92)$$

Lemma 4. For any two rational functions $A(\cdot)$ and $B(\cdot, \cdot)$, the following is true:

$$\begin{aligned} & \text{Re}\{e^{j(\omega_1 k - \varphi_1)}A(e^{j\omega_1 z})\}\text{Re}\{e^{j(\omega_2 k - \varphi_2)}B(z, e^{j\omega_2 z})[v(k)]\} \\ &= \frac{1}{2}\text{Re}\{e^{j((\omega_2 - \omega_1)k + \varphi_1 - \varphi_2)}A(e^{-j\omega_1 z})B(z, e^{j\omega_2 z})[v(k)]\} \\ &+ \frac{1}{2}\text{Re}\{e^{j((\omega_1 + \omega_2)k - \varphi_1 - \varphi_2)}A(e^{j\omega_1 z})B(z, e^{j\omega_2 z})[v(k)]\}. \end{aligned} \quad (93)$$

Lemma 5. For any rational function $B(\cdot, \cdot)$, the following is true:

$$\begin{aligned} & \text{Re}\{e^{j(\omega k - \varphi)}B(z, e^{j\omega z})[v(k)]\} = \cos(\omega k - \varphi)\text{Re}\{B(z, e^{j\omega z})[v(k)]\} \\ & - \sin(\omega k - \varphi)\text{Im}\{B(z, e^{j\omega z})[v(k)]\}. \end{aligned} \quad (94)$$

Lemma 6. Given a transfer function $H(z)$ with its poles inside the unit circle, the signal

$$\cos(\omega k - \varphi)H(z)[v(k)] \quad (95)$$

can be represented in state space form as $(A, B, C(k), D(k))$ where A is exponentially stable.

Proof. Let (A, B, C, D) be the minimal state space realization of $H(z)$ where A is exponentially stable since the poles of $H(z)$ are inside the unit circle. Then, let $x(k)$ be the state vector of $H(z)$ to express $\cos(\omega k - \varphi)H(z)[v(k)]$ in state space form as

$$x(k+1) = Ax(k) + Bv(k), \quad (96)$$

$$\begin{aligned} y(k) &= \cos(\omega k - \varphi)Cx(k) + \cos(\omega k - \varphi)Dv(k), \\ &= C(k)x(k) + D(k)v(k). \quad \square \end{aligned} \quad (97)$$

References

- [1] K.B. Ariyur, M. Krstic, Real-Time Optimization by Extremum-Seeking Control, Wiley-Interscience, Hoboken, NJ, 2003.
- [2] E.-W. Bai, L.-C. Fu, S.S. Sastry, Averaging analysis for discrete time and sampled data adaptive systems, IEEE Transactions on Circuits and Systems 35 (2) (1988) 137–148.
- [3] R. Bhatia, Matrix Analysis, Springer, New York, NY, 1996.
- [4] S. Boyd, L. Vandenberghe, Convex Optimization, Cambridge University Press, New York, NY, 2004.
- [5] D. Carnevale, A. Astolfi, C. Centioli, S. Podda, V. Vitale, L. Zaccarian, A new extremum seeking technique and its application to maximize RF heating on FTU, Fusing Engineering and Design 84 (2009) 554–558.
- [6] Y. Chen, C. Wen, Iterative Learning Control: Convergence, Robustness, and Applications, Springer, New York, NY, 1999.
- [7] J.-Y. Choi, M. Krstic, K.B. Ariyur, J.S. Lee, Extremum seeking control for discrete-time systems, IEEE Transactions on Automatic Control 47 (2002) 318–323.
- [8] J. Cochran, E. Kanso, S.D. Kelly, H. Xiong, M. Krstic, Source seeking for two nonholonomic models of fish locomotion, IEEE Transactions on Robotics 25 (5) (2009) 1166–1176.
- [9] J. Cochran, M. Krstic, Nonholonomic source seeking with tuning of angular velocity, Transactions on Automatic Control 54 (4) (2009) 717–731.
- [10] T.F. Coleman, Y. Li, A reflective Newton method for minimizing a quadratic function subject to bounds on some of the variables, SIAM Journal of Optimization 6 (4) (1996) 1040–1058.
- [11] T. Dierks, B.T. Thumati, S. Jagannathan, Optimal control of unknown affine nonlinear discrete-time systems using offline-trained neural networks with proof of convergence, Neural Networks 22 (2009) 851–860.
- [12] P. Frihauf, M. Krstic, T. Başar, Nash equilibrium seeking with infinitely-many players, in: Proceedings of the American Control Conference, San Francisco, CA, 2011.
- [13] P. Frihauf, M. Krstic, T. Başar, Nash equilibrium seeking in noncooperative games, IEEE Transactions on Automatic Control 57 (5) (2012) 1192–1207.
- [14] W. Fulton, Eigenvalues, invariant factors, highest weights, and Schubert calculus, Bulletin of the American Mathematical Society 37 (3) (2000) 209–249.
- [15] A. Ghaffari, M. Krstic, D. Nešić, Multivariable Newton-based extremum seeking, Automatica 48 (8) (2012) 1759–1767.
- [16] N.J. Killingsworth, S.M. Aceves, D.L. Flowers, F. Espinosa-Loza, M. Krstic, HCCI engine combustion-timing control: optimizing gains and fuel consumption via extremum seeking, IEEE Transactions on Control Systems Technology 17 (6) (2009) 1350–1361.
- [17] J.P. Krieger, M. Krstic, Extremum seeking based on atmospheric turbulence for aircraft endurance, AIAA Journal of Guidance, Control, and Dynamics, 34 (2011) 1876–1885.
- [18] F.L. Lewis, G. Vamvoudakis, Reinforcement learning for partially observable dynamic processes: adaptive dynamic programming using measured output data, IEEE Transactions System Man, and Cybernetics, Part B 41 (1) (2011) 14–25.
- [19] L. Luo, E. Schuster, Mixing enhancement in 2D magnetohydrodynamic channel flow by extremum seeking boundary control, in: Proceedings of the American Control Conference, St. Louis, MO, June 2009.
- [20] K.L. Moore, Iterative Learning Control of Deterministic Systems, Springer-Verlag, New York, NY, 1993.
- [21] S. Ohtake, M. Yamakita, Adaptive output optimal control algorithm for unknown system dynamics based on policy iteration, in: Proceedings of the American Control Conference, Baltimore, MD, June 2010.
- [22] Y. Ou, C. Xu, E. Schuster, T.C. Luce, J.R. Ferron, M.L. Walker, D.A. Humphreys, Design and simulation of extremum-seeking open-loop optimal control of current profile in the DIII-D tokamak, Plasma Physics and Controlled Fusion 50 (11) (2008) 1–24.
- [23] W.J. Rugh, Linear Systems Theory, Prentice Hall, 1996.
- [24] M.S. Stanković, K.H. Johansson, D.M. Stipanović, Distributed seeking of Nash equilibria with applications to mobile sensor networks, IEEE Transactions on Automatic Control 57 (4) (2012) 904–919.
- [25] M.S. Stanković, D.M. Stipanović, Extremum seeking under stochastic noise and applications to mobile sensors, Automatica 46 (2010) 1243–1251.
- [26] Y. Tan, D. Nešić, I. Mareels, On non-local stability properties of extremum seeking control, Automatica 42 (6) (2006) 889–903.
- [27] D. Wang, D. Liu, Q. Wei, D. Zhao, N. Jin, Optimal control of unknown nonaffine nonlinear discrete-time systems based on adaptive dynamic programming, Automatica 48 (8) (2012) 1825–1832.
- [28] H.-H. Wang, S. Yeung, M. Krstic, Experimental application of extremum seeking on an axial-flow compressor, IEEE Transactions on Control Systems Technology 8 (2000) 300–309.
- [29] H.-H. Wang, M. Krstic, Extremum seeking for limit cycle minimization, IEEE Transactions on Automatic Control 45 (2000) 2432–2437.
- [30] J.-X. Xu, Y. Tan, Linear and Nonlinear Iterative Learning Control, Springer-Verlag, 2003.
- [31] Q. Yang, S. Jagannathan, Reinforcement learning controller design for affine nonlinear discrete-time systems using online approximators, IEEE Transactions System Man, and Cybernetics, Part B 42 (2) (2012) 377–390.
- [32] C. Zhang, D. Arnold, N. Ghods, A. Siranosian, M. Krstic, Source seeking with nonholonomic unicycle without position measurement and with tuning of forward velocity, Systems & Control Letters 56 (2007) 245–252.